Group

Tallahassee, FL
4901 Tower Court
Tallahassee, FL 32303
(850) 562-1211
(850) 562-3838 FAX
http://www ersgroup com

APPENDIX A

MARY DUNN BAKER

4901 Tower Court • Tallahassee, FL 32303 • (850) 562-1211
mbaker@ersgroup.com

PROFESSIONAL EXPERIENCE:

ERS Group
- Director (December 1998 to present)
- Vice President and Senior Research Economist (1986 – November 1998)

Design and conduct statistical analyses of economic issues and the valuation of economic losses with emphasis in the area of labor economics. Present expert testimony regarding statistical analysis of employment practices and estimates of economic losses before federal and state courts and in other judicial settings. Develop and implement systems to monitor employment practices. Plan, organize and present seminars on the use of economics and statistics in employment discrimination cases.

Florida State University
- Adjunct Professor of Economics (1988-1991)

Courses taught: Intermediate Microeconomic Theory and Principles of Microeconomics

Auburn University at Montgomery
- Instructor of Economics (1984-1986)

Courses taught: Principles of Economics I and II and Contemporary Economic Problems

Florida Public Service Commission
- Planning and Research Economist, Communications Department (1984)

Analysis of issues in managing the transition to deregulation in the telecommunications industry. Presented recommendations before the Commission.

Florida State University, Department of Economics
- Graduate Instructor (1981-1984)

Courses taught: Economics I and II

Auburn University at Montgomery
- Instructor of Economics (1976-1981)

Courses taught: Concept of Business, Principles of Economics I and II, Labor Economics, Money and Banking, History of Economic Thought and Comparative Economic Systems

EDUCATION:

Ph.D., Florida State University, Economics, 1986

M.S., Auburn University, Economics, 1977

B.A., Auburn University, Political Science/Economics, 1974

Education (Cont.)

HONORS AND AWARDS:

Alpha Lambda Delta (Freshman Women)
Omicron Delta Epsilon (Economics)
Pi Sigma Alpha (Political Science)
Phi Kappa Phi
Alpha Gamma Delta Graduate Scholarship 1981, 1982

SPECIALIZATION:

Labor, Regulated Industries, Natural Resources Economics

OTHER PROFESSIONAL ACTIVITIES:

Consultant to Hazards Management Group, Inc. Developed residential property valuation model for potentially hazardous areas.

PUBLICATIONS AND RESEARCH PAPERS:

Compensation Analysis," Chapter 10 in The Human Resources Program Evaluation Handbook, edited by Jack E. Edwards, John C. Scott and Nambury S. Raju, Sage Publications, 2003.

Estimating Economic Damages in Employment Discrimination Cases," Getting to Verdict: Trying Employment Claims to the Jury, The Florida Bar Labor and Employment Law Section, September 15, 1995.

Economic and Statistical Evidence of Employment Discrimination," The Burden of Proof in Trial of an Employment Case, The Florida Bar Labor and Employment Law Section, September 27, 1991.

Issues Surrounding the Reagan Tax Cuts," Alabama Business and Economic Reports, Center for Business and Economic Research, Auburn University at Montgomery, October 1981, pp. 32-34.

Rising Inflation Reduces Police Pay," (with Ken McCreedy), Alabama Peace Officers Journal, January - February 1981, pp. 18-22.

Social Security - A Summary Outlook," Alabama Business and Economic Reports, Center for Business and Economic Research, Auburn University at Montgomery, July 1980, pp. 2-13.

Social Security and the Decline in Labor Force Participation of Older Men," (Abstract), Journal of Alabama Academy of Science, April 1978, p. 94.

PRESENTATIONS/PROFESSIONAL MEETINGS:

Pay Check Envy: The New Frontier in Employment Litigation," invited speaker at Stearns Weaver Miller's Fifteenth Annual Seminar on Labor & Employment Law, Miami, Florida, May 11, 2005.

Demystifying Compensation Analysis: Analyzing Pay in the Regulatory Environment," invited speaker at the Huntsville Alabama Industry /OFCCP Liaison Group's April 22, 2005 seminar.

Using Statistics to Support and Defend Adverse Impact Age Discrimination Claims," invited speaker at the Regional meeting of the ABA-EEO Committee for Liaison with EEOC, OFCCP and DOJ, New Orleans, Louisiana April 20, 2005.

Presentations/Professional Meetings (Cont.)

Demystifying Compensation Analysis," invited speaker at the South Florida Industry/OFCCP Liaison Group's December 7, 2004 seminar.

Commonly Used Statistical Techniques for Analyzing Selections" and "Analyzing Compensation," invited speaker at the U.S. Department of Labor's Employment Standards Administration Office of Federal Contract Compliance Programs' Western Regions' Training Conference, Tucson, Arizona, November 19, 2003.

The Use and Abuse of Statistics in Employment Discrimination Cases," invited speaker at the Regional Meeting of the ABA-EEO Committee for Liaison with EEOC, OFCCP and DOJ, New Orleans, Louisiana, June 19, 2002.

Preparing For and Responding To OFCCP Salary Audits," invited speaker at the 2001 Aerospace Employment Law Conference, Denver, Colorado, November 29, 2001.

The Use of Statistical Evidence in Employment Discrimination Litigation," invited panel member at the American Bar Association's Section on Labor and Employment Law, Equal Employment Opportunity Committee's Mid-Winter Meeting, Sanibel Island, Florida, March 29, 2001.

Statistical Analyses of Compensation," Lorman Education Services' Seminar, Affirmative Action Compliance in Florida, Tallahassee, Florida, March 16, 2001.

Detecting And Understanding Pay Disparities," invited speaker at the 19th Annual Corporate Counsel Institute, Atlanta, Georgia, December 7, 2000.

"Detecting and Measuring Pay Disparities: Statistical Analyses of Compensation," Economic Research Services seminar on Pay Equity: The New Discrimination Frontier, Atlanta, Georgia, October 27, 2000.

Statistical Analysis of Employment Practices – Selections and Compensation," Lorman Education Services' Seminar, Affirmative Action Compliance In Alabama, Birmingham, Alabama, September 28, 2000.

Pay Equity Analysis in the New Millennium," invited speaker at Jacksonville Industry Liaison Group's Seminar, Business and Government: A Strategic Alliance for 2000, Jacksonville, Florida, May 23, 2000

The Use of Experts in Employment Cases," invited speaker at The University of Alabama School of Law Conference on Employment Law for Government and Public Sector Lawyers, New Orleans, Louisiana, November 1999.

Salary Analyses -- OFCCP and Alternative Methods," O'Melveny & Myers' Seminar, OFCCP Targets Wall Street, New York, New York, February 4, 1999.

Choosing an Age Break: A Discussion of the Issues and an Alternative to Traditional Analyses of Age Discrimination," with Joshua Gotkin, paper presented at the Southern Economic Association Annual Conference, Washington, D.C., 1996.

Estimating Economic Damages in Employment Discrimination Cases," invited speaker at Getting to Verdict: Trying Employment Claims to the Jury, The Florida Bar, Continuing Legal Education Committee and The Labor and Employment Law Section, Orlando, Florida, 1995.

The Use of Statistics in Discrimination Cases," invited speaker at Investing in Human Resources, 1995 Human Resource Development/Personnel Management Conference, Florida Department of Management Services, Tampa, Florida, 1995.

Use and Abuse of Statistics in Employment Claims," invited speaker at Current Issues in Employment Law, Professional Education Systems, Inc., Daytona Beach and Tampa, Florida, 1995.

Methodologies for Determining Economic Damages or Possible Exposure in Employment Discrimination Cases," Economic Research Services, Inc.'s seminar on Economic and Statistical Analysis of Employment Discrimination, 1995-1998.

Presentations/Professional Meetings (Cont.)

Statistical Analyses of Employment Practices in a Recent Innovative Consent Decree," paper presented at the Southern Economic Association Annual Conference, Orlando, Florida, 1994.

Dollars and Sense: Quantifying and Managing Employment Decisions," invited speaker at The American Association of Affirmative Action, Region IV Conference, Savannah, Georgia, 1994.

Commonly Used Statistical Techniques" and "Methods of Analyzing Compensation," ERS Group's seminar on Economic and Statistical Evidence of Employment Discrimination: 1992-2004.

Economic and Statistical Evidence of Employment Discrimination," invited speaker at The Florida Bar Employment Law Litigation Seminar, Orlando, Florida, 1991.

The Current and Future State of the Florida Economy," invited speaker at Tomorrow's Trends Today: An Economic Symposium, Sarasota, Florida, 1988.

The Economic Impact of Fixed Facilities on Residential Communities," invited speaker at the Hazardous Materials Advisory Council Semi-Annual Conference, Orlando, Florida, 1987.

Property Values and Potentially Hazardous Production Facilities," paper presented at American Association of Geographers Annual Meeting, Portland, Oregon, 1987.

PROFESSIONAL ASSOCIATIONS AND MEMBERSHIPS:

American Economic Association

Southern Economic Association

National Association of Forensic Economists

Committee on the Status of Women in the Economics Profession

Auburn University MBA Advisory Board

Auburn University College of Business Advisory Board

Group

APPENDIX A

Tallahassee, FL
4901 Tower Court
Tallahassee, FL 32303
(850) 562-1211
(850) 562-3838 FAX
http://www.ersgroup.com

MARY D. BAKER

4901 Tower Court • Tallahassee, FL 32303 • (850) 562-1211
mbaker@ersgroup.com

DEPOSITION AND COURTROOM TESTIMONY
2001 - 2005

American Association of University Professors v. University of Cincinnati, Gender Pay Equity Dispute, FMCS No. 52 390 001512 00 (2000). [deposition, testimony]

Stephen K. Chen, v. MCI Worldcom; United States District Court, Eastern District of Tennessee at Knoxville, Civil Action No. 3:99-CV-322. [deposition]

John Preuc, et al., v. Continental Micronesia, Inc., and Continental Airlines, Inc.; Superior Court of Guam, Civil Case No. CV69-96. [deposition]

Robert Cram v. Karsten Manufacturing Corporation; United States District Court, District of Arizona, Case No. CIV 00-128 PHX RCB. [deposition]

Frank Minton v. American Bankers Insurance Group, et al.; United States District Court, Southern District of Florida, Miami Division, Case No. 00-3376-CIV-Seitz/Bandstra. [deposition]

Sharon Pollard v. E.I. Du Pont de Nemours Company; United States District Court, Western District of Tennessee, Western Division, Case No. 98-2147 MIA. [deposition]

Curtis Major, et al., v. Eller Media Company; United States District Court, Southern District of Florida, Miami Division, Case No. 00-3870-CIV-MORENO/DUBE. [deposition]

Eugene F. Koren, et al., v. SUPERVALU Inc. and Preferred Products, Inc.; United States District Court, District of Minnesota, Civil No.00-1479 (PAM/JGL). [deposition]

Sylvester McClain, et al., v. Lufkin Industries, Inc.; United States District Court for the Eastern District Court of Texas, Lufkin Division, Civil Action No. 9:97-CV-063. [deposition]

Kimberly Farrow v. Bank of America Corporation; United States District Court, Middle District of Florida, Orlando Division Case No. 6:02-CV-936-OR-28-KRS. [deposition, testimony]

Rhonda Thomas v. Deloitte Consulting, United States District Court, Northern District of Texas, Dallas Division, Case No. 3:02-CV-0343-M. [testimony]

Thaddeus M. Korbin v.Public Service Company of New Mexico; United States District Court for the District of New Mexico, Case No. CV-03-1167 MV/MLT. [deposition]

APPENDIX B

Census 2000 Special EEO Tabulation
EEO-1 Job Categories by Race/Ethnicity, Gender and Geographic Area
Hamilton-Middletown, OH PMSA

Gender	Recorded Labor Force	Total	Caucasian	Minority	African American	Hispanic	Asian or Pacific Is	American Indian	Balance
Officials and Managers (1)									
Total:	19,500	19,492 100.00%	18,170 93.22%	1,322 6.78%	714 3.66%	124 0.64%	379 1.94%	75 0.38%	30 0.15%
Male:		12,618 64.73%	11,885 60.97%	733 3.76%	360 1.85%	69 0.35%	239 1.23%	35 0.18%	30 0.15%
Female:		6,874 35.27%	6,285 32.24%	589 3.02%	354 1.82%	55 0.28%	140 0.72%	40 0.21%	0 0.00%
Professionals (2)									
Total:	31,360	31,350 100.00%	28,425 90.67%	2,925 9.33%	1,265 4.04%	365 1.16%	1,050 3.35%	145 0.46%	100 0.32%
Male:		14,555 46.43%	13,045 41.61%	1,510 4.82%	565 1.80%	175 0.56%	630 2.01%	85 0.27%	55 0.18%
Female:		16,795 53.57%	15,380 49.06%	1,415 4.51%	700 2.23%	190 0.61%	420 1.34%	60 0.19%	45 0.14%
Technicians (3)									
Total:	4,055	4,052 100.00%	3,785 93.41%	267 6.59%	160 3.95%	28 0.69%	75 1.85%	0 0.00%	4 0.10%
Male:		1,548 38.20%	1,440 35.54%	108 2.67%	65 1.60%	8 0.20%	35 0.86%	0 0.00%	0 0.00%
Female:		2,504 61.80%	2,345 57.87%	159 3.92%	95 2.34%	20 0.49%	40 0.99%	0 0.00%	4 0.10%
Sales Workers (4)									
Total:	18,890	18,903 100.00%	17,580 93.00%	1,323 7.00%	710 3.76%	134 0.71%	340 1.80%	90 0.48%	49 0.26%
Male:		9,543 50.48%	8,955 47.37%	588 3.11%	355 1.88%	19 0.10%	130 0.69%	65 0.34%	19 0.10%
Female:		9,360 49.52%	8,625 45.63%	735 3.89%	355 1.88%	115 0.61%	210 1.11%	25 0.13%	30 0.16%
Administrative Support Workers (5)									
Total:	28,790	28,794 100.00%	26,665 92.61%	2,129 7.39%	1,530 5.31%	265 0.92%	195 0.68%	95 0.33%	44 0.15%
Male:		6,679 23.20%	6,250 21.71%	429 1.49%	290 1.01%	90 0.31%	25 0.09%	20 0.07%	4 0.01%
Female:		22,115 76.80%	20,415 70.90%	1,700 5.90%	1,240 4.31%	175 0.61%	170 0.59%	75 0.26%	40 0.14%

APPENDIX B

Census 2000 Special EEO Tabulation
EEO-1 Job Categories by Race/Ethnicity, Gender and Geographic Area
Hamilton-Middletown, OH PMSA

Gender	Recorded Labor Force	Total	Caucasian	Minority	African American	Hispanic	Asian or Pacific Is	American Indian	Balance
Craft Workers (6)									
Total:	16,345	16,347 100.00%	15,130 92.56%	1,217 7.44%	575 3.52%	384 2.35%	75 0.46%	169 1.03%	14 0.09%
Male:		15,480 94.70%	14,355 87.81%	1,125 6.88%	515 3.15%	380 2.32%	55 0.34%	165 1.01%	10 0.06%
Female:		867 5.30%	775 4.74%	92 0.56%	60 0.37%	4 0.02%	20 0.12%	4 0.02%	4 0.02%
Operatives (7)									
Total:	22,120	22,114 100.00%	19,790 89.49%	2,324 10.51%	1,410 6.38%	340 1.54%	335 1.51%	170 0.77%	69 0.31%
Male:		16,344 73.91%	14,760 66.75%	1,584 7.16%	965 4.36%	225 1.02%	210 0.95%	130 0.59%	54 0.24%
Female:		5,770 26.09%	5,030 22.75%	740 3.35%	445 2.01%	115 0.52%	125 0.57%	40 0.18%	15 0.07%
Laborers and Helpers (8)									
Total:	6,845	6,848 100.00%	6,110 89.22%	738 10.78%	280 4.09%	304 4.44%	45 0.66%	90 1.31%	19 0.28%
Male:		5,924 86.51%	5,250 76.66%	674 9.84%	260 3.80%	300 4.38%	20 0.29%	90 1.31%	4 0.06%
Female:		924 13.49%	860 12.56%	64 0.93%	20 0.29%	4 0.06%	25 0.37%	0 0.00%	15 0.22%
Service Workers (9)									
Total:	21,985	21,985 100.00%	19,460 88.51%	2,525 11.49%	1,640 7.46%	295 1.34%	325 1.48%	190 0.86%	75 0.34%
Male:		8,085 36.78%	7,150 32.52%	935 4.25%	620 2.82%	100 0.45%	125 0.57%	65 0.30%	25 0.11%
Female:		13,900 63.22%	12,310 55.99%	1,590 7.23%	1,020 4.64%	195 0.89%	200 0.91%	125 0.57%	50 0.23%
Unemployed, No Civilian Work Experience Since 1995 (10)									
Total:	470	467 100.00%	385 82.44%	82 17.56%	60 12.85%	14 3.00%	4 0.86%	0 0.00%	4 0.86%
Male:		270 57.82%	210 44.97%	60 12.85%	50 10.71%	10 2.14%	0 0.00%	0 0.00%	0 0.00%
Female:		197 42.18%	175 37.47%	22 4.71%	10 2.14%	4 0.86%	4 0.86%	0 0.00%	4 0.86%

APPENDIX B

Census 2000 Special EEO Tabulation
EEO-1 Job Categories by Race/Ethnicity, Gender and Geographic Area
Hamilton-Middletown, OH PMSA

Gender	Recorded Labor Force	Total	Caucasian	Minority	African American	Hispanic	Asian or Pacific Is	American Indian	Balance
Total Civilian Labor Force									
Total:	170,360	170,352	155,500	14,852	8,344	2,253	2,823	1,024	408
		100.00%	*91.28%*	*8.72%*	*4.90%*	*1.32%*	*1.66%*	*0.60%*	*0.24%*
Male:		91,046	83,300	7,746	4,045	1,376	1,469	655	201
		53.45%	*48.90%*	*4.55%*	*2.37%*	*0.81%*	*0.86%*	*0.38%*	*0.12%*
Female:		79,306	72,200	7,106	4,299	877	1,354	369	207
		46.55%	*42.38%*	*4.17%*	*2.52%*	*0.51%*	*0.79%*	*0.22%*	*0.12%*

Notes

1. The total Civilian Labor Force shown for each Census Detailed Occupation category may differ from the Total Count due to the rounding scheme employed by the Census Bureau in construction of the file to protect individual's privacy. For details see "Appendix 5: Protecting Privacy, Census 2000 Special Tabulation Technical Documentation".

2. The race/ethnicity groupings are constructed according to guidance provided by the EEOC in, "Introduction to Race and Ethnic (Hispanic Origin) Data for the Census 2000 Special EEO File".

Caucasian: White not Hispanic or Latino.

Minority: Total minus White not Hispanic or Latino.

African American: Black or African American not Hispanic or Latino, plus Black or African American and White not Hispanic.

Hispanic: White Hispanic or Lation plus Other Hispanic or Latino.

Asian: Asian not Hispanic or Lation plus Asian and White not Hispanic or Latino.

American Indian: American Indian or Alaska Native not Hispanic or Latino, plus American Indian or Alaska Native and White not Hispanic or Latino.

Pacific Islander: Native Hawaiian or Other Pacific Islander not Hispanic or Latino.

Balance: American Indian or Alaska Native and Black or African American not Hispanic or Latino, plus balance of individuals reporting more than one race not Hispanic or Latino, plus individuals reporting some other race not Hispanic or Latino.

Source: Census 2000 Special EEO Tabulation, File 4: Civilian Labor Force 16 years and older: EEO-1 Job Categories by Race/Ethnicity and Sex. Sponsored by Equal Employment Opportunity Commission, Department of Justice, Department of Labor, Office of Federal Contract Compliance Program and Office of Personnel Management. Prepared by U.S. Census Bureau. Issued February, 2004.

Appendix C
Percent Black Among Laborers and Helpers (EEO-1 Category 8)

State	County	Percent of Ashland's Current Union Employees Who Reside in County, State	Total Number of Laborers & Helpers	Number of Black Laborers & Helpers	Percent Black Among Laborers & Helpers
KY	BOYD	32.66%	968	14	1.42%
	CARTER	6.19%	999	0	0.00%
	FLEMING	0.10%	447	34	7.61%
	GREENUP	26.27%	823	8	0.97%
	JOHNSON	0.30%	609	0	0.00%
	LAWRENCE	1.93%	408	0	0.00%
	LEWIS	1.01%	569	0	0.00%
	ROWAN	0.10%	650	0	0.00%
OH	ADAMS	0.10%	911	0	0.00%
	GALLIA	0.10%	712	30	4.21%
	LAWRENCE	20.39%	1365	15	1.10%
	SCIOTO	3.45%	1842	14	0.76%
WV	CABELL	2.64%	1890	110	5.82%
	LINCOLN	0.61%	639	0	0.00%
	MINGO	0.10%	417	0	0.00%
	PUTNAM	0.20%	1039	4	0.39%
	WAYNE	3.85%	928	0	0.00%

Source: Census 2000 Special EEO Tabulation, File 4: Civilian Labor Force 16 years and older: EEO-1 Job Categories by Race/Ethnicity and Sex. Sponsored by Equal Employment Opportunity Commission, Department of Justice, Department of Labor, Office of Federal Contract Compliance Program and Office of Personnel Management. Prepared by U.S. Census Bureau. Issued February, 2004.

APPENDIX D

USE OF STATISTICS IN EQUAL EMPLOYMENT OPPORTUNITY LITIGATION

Walter B. Connolly, Jr.
David W. Peterson
Michael J. Connolly

Updated by
Scott Zesch
Member of the Texas and New Mexico Bars

2004
(Date originally published: 1980)

Law Journal Press
105 Madison Ave.
New York, New York 10016
www.lawcatalog.com

#00553
(Rel. 20)

APPENDIX D

VISIT OUR WEBSITE AT
www.lawcatalog.com

Library of Congress Cataloging-in-Publication Data

ISBN 1-58852-006-4

Connolly, Walter B.,
Use of statistics in equal employment opportunity litigation.
1. Discrimination in employment—Law and legislation—United States. 2. Actions and defenses—United States. 3. Discrimination in employment—United States—Statistical methods. 4. Statistics. I. Peterson, David W., 1940-joint author. II. Title. 3
KF3464.C67 344.73'01133O269' 79-28419

APPENDIX D

§ 11.11 One Tail or Two? Or Does It Really Matter?

[1]—Introduction

The District of Columbia Circuit in *Palmer v. Schultz*[1] wrote one of the most lucid descriptions extant of the statistical standards of proof a court expects from a plaintiff. The court begins its discussion with a description of the statistical method for testing hypotheses, and from there weaves in the concepts of one- and two-tailed tests and their associated probability or p-values. It goes on to assert that "a court should generally adopt a two-tailed approach to evaluating the probability that the contested disparity resulted by chance" and that "statistical evidence must meet the 5% level . . . for it alone to establish a prima facie case. . . ."[2] The court then asserts that "a two-tailed test and a 5% probability of randomness require statistical evidence measuring 1.96 standard deviations" and that "if plaintiffs come into court relying *only* on evidence that the under selection of women for a particular job measures 1.75 standard deviations, it seems improper for a court to establish an inference of disparate treatment on the basis of this evidence alone."[3]

This clearest-yet statement of a statistical standard of proof is a great service to litigants, enabling them better to evaluate their positions before trial. But the wording of the standard seems to imply that one-tailed tests are not proper vehicles for setting forth a statistical case, and perhaps even to imply that all statistical test results should be expressed on the standard deviation scale for ready comparison to the 1.96 threshold. This is unfortunate, and almost surely unintended. It seems likely that the true intent of the court was to hold that a two-tailed test should be interpreted against a 5% probability threshold, that a one-tailed test should be interpreted against a threshold of 2.5%, and, when used with these interpretations, the tests are equally

---

[1] Palmer v. Schultz, 43 F.E.P. Cases 452 (D.C. Cir. 1987).

[2] *Id.* at 465.

[3] *Id.* See also:

Sixth Circuit: Dobbs-Weinstein v. Vanderbilt University, 1 F. Supp. 2d 783, 808 (M.D. Tenn. 1998), *aff'd* 185 F.3d 542 (6th Cir. 1999) (noting that a one-tailed test makes it easier to claim statistically significant evidence of discrimination, since a probability level of 0.5 under a two-tailed analysis translates into 1.96 standard deviations, while under a one-tailed analysis it becomes just 1.65 standard deviations).

District of Columbia Circuit: Moore v. Summers, 113 F. Supp. 2d 5, 20 n.2 (D.D.C. 2000) (since purpose of Title VII is to assure that people are treated equally, not that one group is treated at least as well as or better than another, District of Columbia Circuit prefers two-tailed tests).

satisfactory. It also seems likely that, given the opportunity, the court would make it clearer that its 1.96 standard deviation threshold applies only to certain kinds of random variables.

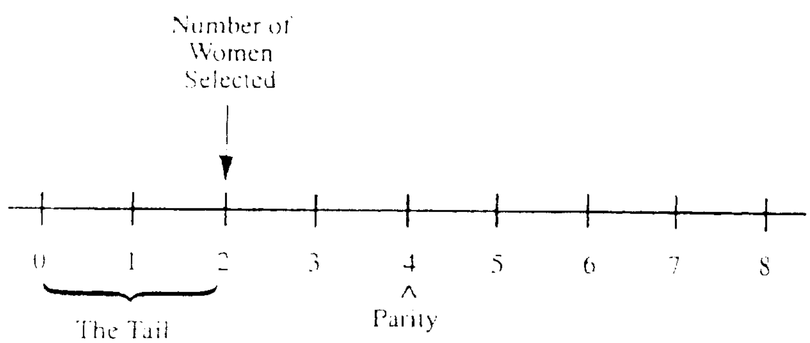
[2]—Statistical Theory of One- and Two-Tailed Tests

There are two commonly used ways to express the results of the test of a statistical hypothesis. The first of these uses probabilities, and the second uses standard deviations. The first of these two methods itself has two variations: the one-tailed probability and the two-tailed probability. The key distinction is that one-tailed analysis measures whether a particular group is disfavored, while two-tailed analysis tests whether the group is either preferred or disfavored.[3.1] The salient features of these methods can be conveyed by examples in which a group of people has been chosen from a much larger group called the pool, and the gender composition of the group selected is contrasted with that of the pool to determine whether the process appears to have favored the selection of men to the disadvantage of women.

[a]—The One-Tailed Probability

The one-tailed probability is the probability that random selection from the large pool would result in the group having as few or fewer

Figure 11.3 The One-Tailed Test



[3.1] *Sixth Circuit:* *Dobbs-Weinstein v. Vanderbilt University*, N. 3 *supra*, 1 F. Supp. 2d at 808 (finding that one tailed analysis was not appropriate in wage discrimination case).

District of Columbia Circuit: Hartman v. Duffey, 88 F.3d 1232, 1238 (D.C. Cir. 1996); Moore v. Summers, 113 F. Supp. 2d 5, 20 n.2 (D.D.C. 2000) (explaining that one-tailed test reveals whether blacks are treated at least as well as or better than

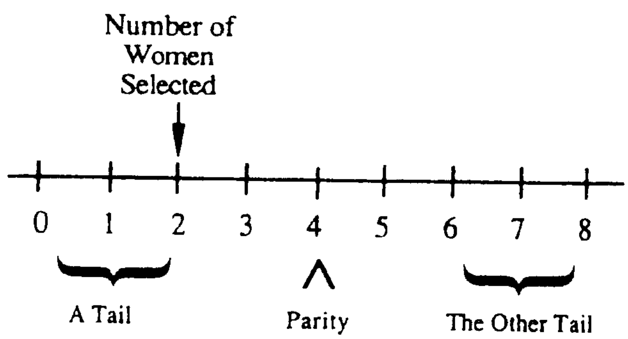
women that it in fact contains. For example, suppose the group contains 8 people, of whom 2 are women, and suppose half of the members of the pool are women. The one-tailed probability is the probability that if 8 people were drawn randomly from the pool, no more than 2 of them would be women. The notion of a tail (see Figure 11.3) arises from the observation that this probability is the sum of the three individual probabilities that the pool might contain, respectively, 2, 1 and 0 women, and that these probabilities pertain to the most extreme (the tail) of the possible outcomes of a random draw. The usefulness of the one-tailed probability (as well as the two-tailed probability) as a measure of imbalance derives from the fact that a small probability value indicates an imbalance of so great a magnitude that it cannot reasonably be attributed to the luck of the draw.

---

whites, but does not indicate whether whites are treated worse than blacks; two-tailed test demonstrates whether blacks and whites were treated equally, taking into account both whether whites are treated as well as or better than blacks and vice versa).

[b]—The Two-Tailed Probability. The two-tailed probability is the probability that if the group were drawn randomly from the pool, the composition of the group would differ *in any sense* by as much or more than it in fact does from that of the pool. Continuing the above example, the two-tailed probability is the probability that if 8 people were drawn randomly from the pool, the number of women among them would be 2 or less or else 6 or more (see Figure 11.4). The underlying rationale is that since half of the pool members are women, a random draw of 8 people is as likely to yield 3 women as 5; to yield 2 women as 6; 1 woman as 7; and so forth. Hence the group's imbalance of 2 women in 8 is just as extreme as an imbalance of 6 women in 8, and less extreme than the imbalance in groups containing 7 or 8 women. By this mode of thought, there are two types of extemes (hence two tails) to consider: the possibilities that a random draw would yield 2, 1 or 0 women, and the possibilities that it would yield 6, 7 or 8 women.

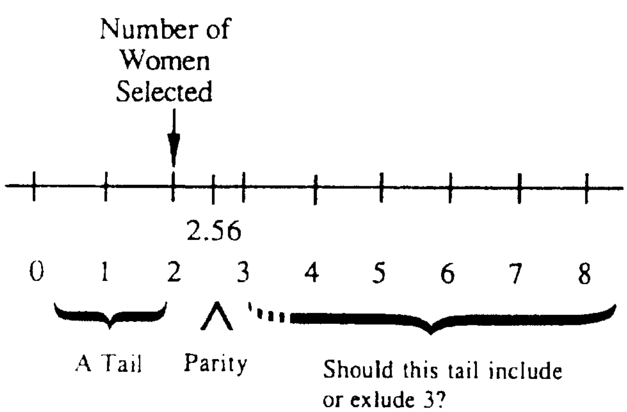
Figure 11.4 The Two-Tailed Test



It is clear from the symmetry of this example that the two-tailed probability is exactly twice the numerical value of the one-tailed probability. Hence whether one reports the degree of imbalance in the example as a one-tailed probability makes little difference, because the recipient of the information can easily convert from one to the other. That is, neither measure of imbalance contains either more or less information than does the other. It is important, of course, to indicate the number of tails on which the reported probability is based.

In the absence of such symmetry, the definition of the two-tailed test may be ambiguous. Suppose for example that 32% (instead of 50%) of the pool members are women. The one-tailed probability[4] (.501) is still the probability that 2 or fewer of the 8 people selected randomly would be women, but it is not clear how one should define the antipodal tail. Perfect parity is 32% of 8, or 2.56 women out of 8. The difference between 2.56 and the number (2) of women in the group is 0.56 women, so perhaps the antipodal threshold should be 2.56 + .56 = 3.12 women. Hence an imbalance of 3.12 or more women (that is, 4 or more) would be considered comparable to an imbalance of 2 or fewer women. By this definition, the two-tailed probability is 0.733. Alternatively, we might note that the probability (0.499) of selecting 3 or more women is nearly the same as (but slightly less than) the probability (0.501) of selecting 2 or fewer women, and on this basis regard an imbalance of 3 or more women to be comparable to an imbalance of 2 or fewer women. By this definition, the two-tailed probability is 0.499 + 0.501 = 1.00, not 0.733 (see Figure 11.5).

Figure 11.5 The One-Tailed Test: Asymmetric Case



This illustrates that different criteria exist for the definition of a two-tailed probability, and that they can lead to different numerical values for the two-tailed probability. In view of this, it is also apparent that a two-tailed probability need not be equal to twice the one-tailed probability.

---

[4] The calculation of this probability and others to follow is somewhat involved, and the details are not presented here. The calculations are based on the well-known binomial statistical model.

APPENDIX D

[c]—The Standard Deviation Scale. The standard deviation method for expressing disparities is an alternative to the probability methods. Unlike probabilites, which take values between 0 and 1 (or 0% and 100%), a disparity measured on the standard deviation scale can be any positive or negative number. In the example of the selection of 8 people from a pool containing equal numbers of men and wome, the group having 2 or more women among the 8 is a certain number of standard devitions away from perfectly reflecting the gender composition of the pool. There is no sense in which this disparity is one-or two-tailed; one simply states that having 2 women in the group of 8 is a certain number of standard deviations below the perfect parity figure of 4.[5]

[d]—Discussion. As the above examples show, a given disparity can sometimes be expressed as a one-tailed probability, a two-tailed probability, and a number of standard deviations. Unfortunately, there is no formula by means of which one can always convert among these measures, because their relationships depend on the specific nature of the probabilities governing the slection process, as well as the definition chosen for the two-tailed probability. However, if the selection probabilities are such that the number of women selected is a so-called normal random variable, then a standard conversion holds: a two-tailed probability is twice the corresponding one-tailed probability, a two-tailed probability of 5% corresponds to a 1.96 standard deviation disparity, and (therefore) a one-tailed probability of 2.5% also corresponds to a 1.96 standard deviation disparity. If the number of women drawn is not a normal random variable (and it is not in any of the examples above), then these relationships need not and often do not hold.

Statisticians generally prefer to use the probability measures of disparity because they are directly comparable across a wide variety of situations. In contrast, a disparity of, say, two standard devia-

---

[5] The standard deviation in this case is calculated using a binomial model:

$$\text{standard deviation} = \sqrt{(8\text{x}.5\text{x}.5)} = 1.414.$$

Hence the number of standard deviations by which 2 differs from 4 is (4 - 2)/1.414 = 1.414.

tions can in some circumstances be rather improbable as a consequence only of chance, and in others, highly probable.[6]

[3]—Discrimination Litigation Conventions.

In many of the published opinions regarding the use of statistics in discrimination litigation the attention of the court has been focused on a random variable that is normal, or nearly so.[7] Consequently, it has made little difference whether the court couched its opinion in terms of one- or two-tailed probabilities, or in numbers of standard deviations, because as long as the units of denomination were clear, the recipient of the information could freely convert to his or her preferred scale.

Indeed, the court of appeals in *Palmer v. Schultz* shifted freely among all three measures in its discussion. But when it fixed the minimal probability threshold (5%) it thought most appropriate, it happened to be thinking in terms of two-tailed probabilities, and it failed to emphasize that the one-tailed equivalent of that threshold is 2.5%.[8] Likewise, when it translated the 5% two-tailed probability threshold into the 1.96 standard deviation threshold, it failed to emphasize that this translation only holds for certain kinds of random variables, most notably the normal random variables.[9] Conse-

---

[6] For example, if a single draw from a pool of people of whom 80% are women results in the selection of a man, a disparity of two standard deviations exists between the composition of the group selected (just one person) and that of the pool. But there is evidently one chance in five (a probability of 20%) of this imbalance occurring when the selection is random. This is substantially larger than the 2.3% probability (one chance in 44) to which two standard deviations corresponds for a normal random variable. Or consider a third situation: A random variable that can take any value between 0 and 1 with equal probability has an average or perfect parity value of .5, and a standard deviation of .289. The greatest amount by which this variable can differ from its perfect parity value is (1–.5)/.289 = 1.73 standard deviations. Hence it is impossible (probability = 0%) for this random variable to register a value as large as two on the standard deviation scale.

[7] See Castaneda v. Partida, 430 U.S. 482, 97 S.Ct. 1272, 51 L.Ed.2d 498 (1977); Hazelwood School District v. United States, 433 U.S. 299, 307, 97 S.Ct. 2736, 53 L.Ed.2d 768 (1977); Palmer v. Schultz, 43 F.E.P. Cases at 460 (D.C. Cir. 1987).

[8] At 43 F.E.P. Cases at 465, n.9, the court does clearly acknowledge the equivalence of a one-tailed probability of 2.5% and a two-tailed probability of 5%. Its failure to mention the equivalence explicitly in its main statement of the standard for a statistical *prima facie* case therefore seems due to its choice of organization for its written opinion, rather than its intent to favor two-tailed probabilities over one-tailed probabilities.

[9] Again, in n.9, *Id.*, the court explicitly recognizes the equivalence of 1.96 standard

quenlty, the reader of *Palmer v. Schultz* is likely to conclude (1) that two-tailed probabilities are always more appropriate than one-tailed probabilities, and (2) that regardless of the underlying probability model, 1.96 standard deviations is an appropriate minimal threshold against which to judge a disparity on the standard deviation scale. A more logical interpretation of this court's view is that it matters not whether a litigant uses a one- or a two-tailed probability, so long as it is used in conjunciton with the appropriate threshold probability: 5% for two-tailed probabilities, and 2.5% for one-tailed probabilities. Furthermore, as long as one is dealing with a normal random variable, one may use a third (completely equivalent) alternative the standard deviation scale along with a threshold of 1.96 standard deviations.

But if one is not dealing with a normal random variable, is the 1.96 standard deviation threshold still appropriate? Not necessarily. The 1.96 number arises in the context of normal random variables *because* it is equivalent to a one-tailed probability of 2.5% and to a two-tailed probability of 5%, and it is not of itself a particularly notworthy or appropriate threshold. For non-normal random variables, therefore, it seems most appropriate to recalibrate the standard deviation scale to determine the threshold which corresponds either to a one-tailed probability of 2.5% or, in cases in which there is no material ambiguity in its definition, to a two-tailed probability of 5%.

Under this interpretation of the intent of the court of appeals in *Palmer*, litigants are free to use one-tailed probabilities, two-tailed probabilities, or standard deviation tests, whichever seem best to fit the fact situation, as long as these various measures of disparity are measured against the proper thresholds for significance.

---

deviations with both a one-tailed probability of 2.5% and a two-tailed probability of 5%. Since the court's discussion is confined to normal random variables (see 43 F.E.P. Cases at 462), however, it is never compelled to note that for other kinds of random variables, 1.96 standard deviations may not even approximate a one-tailed probability of 2.5% or a two-tailed probability of 5%.

APPENDIX E

The Statistics of Discrimination

Using Statistical Evidence in Discrimination Cases

By

RAMONA L. PAETZOLD
Texas A&M University

and

STEVEN L. WILLBORN
University of Nebraska

CBC

CLARK
BOARDMAN
CALLAGHAN®

DEERFIELD, IL · NEW YORK, NY · ROCHESTER, NY

Customer Service: 1-800-323-1336

 Pub 12/96

APPENDIX E

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48–1984.

Library of Congress Cataloging in Publication Data

Paetzold, Ramona L.
The statistics of discrimination : using statistical evidence in discrimination cases / Ramona L. Paetzold & Steven L. Willborn.
p. cm.—(Individual rights series)
Includes index.
ISBN 0–07–172474–5
1. Discrimination—Law and legislation—United States.
2. Actions and defenses—United States. 3. Forensic statistics—United States. 4. Discrimination—United States—Statistical methods.
I. Willborn, Steven L. II. Title. III. Series.
KF8925.D5P34 1994
344.73'01133—dc20
[347.3041133] 94–26042
CIP

Ω

This latter example illustrates a constraint on avoiding Type I errors. By adopting a standard that minimizes the number of Type I errors, the chance of identifying discrimination when it actually occurs is reduced. That is, the possibility that we would fail to reject the null hypothesis even though it was in fact not true is increased. This type of error, failing to reject X when in fact it is not true, is called a *Type II error*.[22] In our last example, using the 1 percent standard for Type I error, a court would not infer that Employer A had discriminated even though, if Employer A had truly not discriminated, one would see this hiring outcome (10 women hired) in fewer than one case out of 20. Thus, the possibility of Type II errors is a constraint on our ability to set strict standards to avoid Type I errors. In the context of discrimination law, for example, we want both (1) to avoid inferences of discrimination when, in fact, the employer did not discriminate (Type I errors) and yet, (2) to avoid failing to make an inference of discrimination when, in fact, the employer did discriminate (Type II errors). Since the goals are in conflict, pursuit of either goal limits pursuit of the other.[23]

§2.05 The Problem of Causation

Causation is particularly important in discrimination cases and requires careful delineation of statistical and legal meanings.[24] In discrimination cases, the ultimate concern is with what causes disparities between two groups. In our systemic disparate treatment example, for instance, liability depends on the cause of the disparity between the proportion of women hired by the employer and the proportion of women in the applicant pool. Was the disparity caused by the employer's intention to discriminate (in which case the employer would be liable)? Or was it caused by other factors?[25]

---

[22] The relationship between Type I and Type II errors is complex, but, if all other factors are held constant, decreasing the risk of Type I errors increases the risk of Type II errors and vice versa. *See* Michael O. Finkelstein & Bruce Levin, Statistics for Lawyers 186-188 (1990). *See also* the discussion of statistical power in §4.15.

[23] A common way to reduce the possibility of both types of error is to increase the sample size. In most discrimination cases, that option is not available, so the conflict becomes more salient.

[24] For a discussion of other tensions between statistical and legal meanings, *see* §2.06.

[25] Causation issues are also important in disparate impact cases. Consider, for example, a case in which a test appears to screen out a higher proportion of women than men. The first issue will be whether the test actually does cause the apparent disparity between male and female pass rates. If the difference in pass rates is large and based on a large sample, statistical analysis may indicate that the difference would have been quite unlikely to occur by chance. Thus, an inference that the

For two basic reasons, however, statistical inference can never *prove*[26] a cause for a disparity.[27] First, statistical inference cannot prove the cause of a disparity, because it provides information only on the likelihood of particular outcomes. In our example, for instance, statistical inference provided information about how often one would observe 10 women hired if the employer repeatedly hired 100 people randomly from an applicant pool consisting of 20 percent women. Statistical analysis indicated only—no more, no less—that one would rarely see as few as 10 women hired under those circumstances. It did not prove that discrimination caused the outcome; the employer could have hired without discrimination and been that one case out of 20 in which as few as 10 women were hired. Indeed, even if the employer had hired no women, statistical analysis would not prove that the employer discriminated; the employer could have been that extremely rare case in which random hiring resulted in no women hired. A fact-finder may (and should) be more willing to infer that discrimination caused a disparity as the employer's hiring outcome becomes less and less likely to have occurred by chance, but the inference of causation is the fact-finder's. Statistical analysis only provides information on how likely it is that the outcome can be explained by other factors, such as chance.[28]

---

test truly does have an adverse impact on women may be appropriate. On the other hand, if the difference in pass rates is small and/or based on a small sample, statistical analysis may indicate that the apparent disparity very likely was caused by chance. As a result, a fact-finder may be unwilling to infer that the test truly has an adverse impact on women. In either case, of course, the causation issue of interest, whether the test actually causes a difference in male and female pass rates, cannot be proven or disproven by statistical inference. Rather, statistical inference can be used only to evaluate the likelihood that the disparity occurred by chance. For detailed discussion, *see* §§5.05-5.10.

[26] Here, of course, "prove" is used to mean "establish conclusively" in the sense in which a mathematician or logician might use the word when verifying results. "Prove" is not used in the legal sense of "sufficient to justify a particular result." The courts have recognized time and time again that statistical methods are capable of proving discrimination in the legal sense.

[27] Some readers may recall that David Hume made quite a name for himself on a stronger, but closely analogous point:

> [W]e never can, by our utmost scrutiny, discover any thing but one event following another, without being able to comprehend any force or power by which the cause operates, or any connexion between it and its supposed effect.

David Hume, Enquiries Concerning the Human Understanding and Concerning the Principles of Morals 73-74 (LA Selby-Bigge ed 2d ed 1902 Clarendon Press) (1777).

[28] Our example focused on chance as the explanatory variable, other than discrimination, which might explain the disparity. Later in this volume, techniques that can

APPENDIX E

Statistical analysis cannot prove the cause of a disparity for another reason as well: statistics cannot eliminate the possibility that factors other than those considered by the analysis may have caused the disparity. In our example, statistical analysis indicated that it was unlikely that chance caused the disparity between the proportion of women hired and the proportion in the applicant pool. One remaining possibility is that discrimination caused the disparity, but it is only one of the remaining possibilities. The employer might claim, for example, that it only hired people who possessed a certain legitimate job qualification, X, and that a higher proportion of male applicants than female applicants possessed X. If the employer's claim is true, rejecting chance as a likely cause of the disparity may not be sufficient to justify an inference of discrimination. The disparity may have been caused by X rather than sex discrimination. The plaintiffs could certainly use statistical techniques (or other types of evidence) to evaluate this employer claim as well, but even if they successfully demonstrate that it is unlikely the disparity was caused by X *and* chance, an infinity of other possible causes for the disparity exists. A fact-finder's inference of discrimination becomes stronger as the statistical evidence (or other evidence) renders unlikely more and more of the plausible alternative explanations for the disparity. But no analysis can consider all of the alternative explanations. As a result, because the logic of statistical analysis operates negatively by rejecting possible causative factors, statistical analysis can never affirmatively prove discrimination.

These formal limitations on the ability of statistical analysis to prove causation are mentioned because they highlight the underlying logic of statistical analysis. At the same time, however, legal proof is considerably different from the rigorous proof required by mathematics or formal logic. The courts have clearly and correctly recognized that statistical analysis alone can be sufficient to prove causation in a legal sense[29] despite its limitations in proving causation in a more rigorous sense. And the courts have recognized that the statistical analysis itself need not be perfect to have these legal consequences.[30] At the end of the day, of course, the relevant standard for evaluating statistical methods is not whether they meet the standards of mathematics or formal logic, but whether they meet the requirements of legal proof.[31]

---

be used to analyze the likelihood that other variables, such as job tenure or qualifications, resulted in particular outcomes will be discussed. *See* chs 4, 5, and 6.

[29] *See* Teamsters v United States, 431 US 324, 339 (1977).

[30] Bazemore v Friday, 478 US 385 (1986).

[31] In a very real sense, the requirements of legal proof are what this book is about: What degree of statistical rigor is required of plaintiffs when establishing a prima facie case through statistical evidence? What is required of defendants who object to the statistical analyses of plaintiffs or who forward their own statistical analyses? What weight should courts place on imperfect statistical analyses?